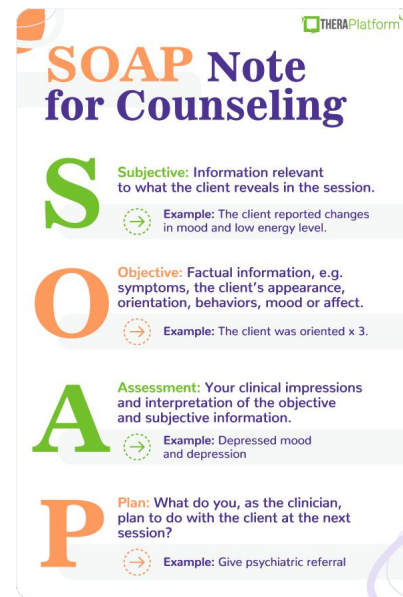


Maintaining Consistency in extended Multiple Text generation

Aadit Deshpande

Background - Coherence over Longer Texts

- Initial Objective - Generate synthetic SOAP notes
 - Privacy concerns + Synthetic data required for downstream IR tasks
- SOAP notes
 - Subjective, Objective, Assessment, Plan
- Requirements
 - Sufficiently detailed and factually correct (independently)
 - Info in each section be consistent with each other!
 - That the 4 sections be related to each other
 - That we can parametrize the generation of these sections
- **Consistency among Longer Texts (4 turns)**
- Approaches
 - One model to generate everything (problematic!)
 - Generate each section independently (inconsistency issues)



Motivation - Consistency in Multiple Shorter Texts

- How do we maintain consistency among “turn-wise” text generation
 - Natural Extensions => Control text length, Control # turns!
- What kinds of “Multiple Texts”?
 - Create Personas with **Medical Profiles + Personal Profiles**
 - **Blog Posts**
 - More personal and expressive,
 - lesser constraints on content/format
 - Chances to discover model errors
- What is “Consistency” exactly?
 - **Coherence** - Factual?
 - **Plausibility** - Consistent w/ estd facts?
 - **Time-Scale** - How many turns is it coherent?



Architecture - What Worked

- Architecture
- **Cascaded Summarizer + Generator**
- Models
 - Generator - Llama-2-7b-chat-hf
 - Summarizer
 - Llama-2-7b-chat-hf (One-shot with task description)
 - **Maximal Marginal Relevance (MMR)**
 - *(used in diversifying search engine rankings)*
- Past-look Over - Best seems to be 2 turns!

- Problem Setup
- **Universal Fact List** - Parametrize!!
 - Condition
 - Severity
 - Age
 - Gender
 - Profession
 - Location
 - Medications
 - Activity Restrictions
 - Allergens
 - Food/Substance Restrictions
- Prompt Structure
 - System Prompt + Task Prompt + Summary Prompt

Architecture

Universal Fact List

1: ["Asthma is a chronic lung disease caused by inflammation and muscle tightening around the airways, which makes it harder to breathe", **# CONDITION**

"Your symptoms include coughing, wheezing, shortness of breath and chest tightness and range from mild to severe", **# SEVERITY**

"You use a Pulmicort Flexhaler and Perforomist inhalation solution to mitigate your asthma symptoms", **# MEDICATION**

"You can partake in intense physical activity but not for more than 45 minutes a day", **# ACTIVITY**

"You can not participate in cold-weather sports like ice hockey, skiing or ice skating", **# ACTIVITY**

"You live in Pasadena, California", **# LOCATION**

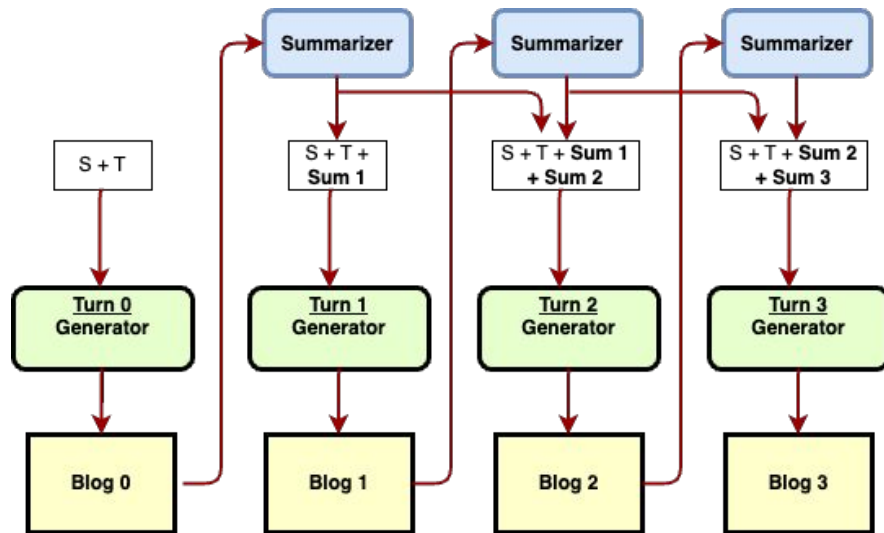
"You have severe pet allergies which trigger your asthma",

ALLERGEN

"Strong alcoholic drinks trigger your asthma symptoms"]],

SUBSTANCE

Pipeline



Prompt Example

System Prompt:

You are a 30 year-old woman with asthma. You work as an actor, writing your monthly blog based on the interesting events in your life. Write about how the following facts change over time: 1. Aaa 2. Bbb 3.Ccc 4. Ddd

Summary Prompt:

This is a summary of all your previous blog posts: February has been quite a challenging month for me so far due to various reasons One major issue ha been increased incidence of allergic reactions online in recent week Sneezing fits coupled with itching n spells of wheezing have become my new nightmare These attacks are particularly difficult because they often occur when least expect them , making it hard to provide proper medication and avoid potential triggers The other significant setback is dealing with the weather changes. This being California where the sun shine bright almost everyday, its challengings to maintain adequate drink intake when the outside temperature drops below 60 degrees Fahrenheit It's Hard to maintain consistent relief from steroid withdrawal Title: A New Season Brings Changes (And Challenge)It feels like just yesterday w e were welcoming spring into their new year(or at least what felt like new) Now here we are in the eighth month of the year and our bodily systems r already signalling changes ahead however these rising heat indexes give birth too greater concerns regarding how my body reacts to stimuli A lo ta weeks now,my upper respiratory system has been protesting a possible allergy attack, prompting me to toss aside my regular routine in favor of quick relief meds

Task Prompt:

Now write a blog post for the month of April Blog Post:

Architecture - What did NOT work :(

- Fact-checking Models
 - 2 NLI Models
 - BART-large-mnli
 - mDeBERTa-v3-base-xnli-multilingual-nli
 - Heavily biased towards Neutral option
 - Struggles with numericals
- Summarizer models
 - FlanT5-xl-3B, BART-Large-CNN
 - FlanT5 - Brevity is a big issue!
 - BART - Defaulted to first three sentences...not good for blog use case!
- Meditron-7b (Continued pre-training)
 - Research-paper style writing
 - Stylistic inconsistency!

Example

Title: A new beginning! As I welcome another winter season at my doorstep, can't help but reflect back on last years progress towards managing my condition well into this one..and boy oh boy there have been some changes that definitely deserve an update here folks because let me tell ya it's not just those same old resolutions we make every December (even though they still hold true). First off remember point number two from our September diatribe? That "chronic lung disease known as Asthma" has indeed taken center stage once again during these coldest months ahead - so much so sometimes even after using both medics prescribed medications plus natural remedies like deep breathing exercises whenever possible – all while keeping close tabs monitor daily peak flow readings or PEF levels found within those handy dandy little gadgets attached near my bedside nightstand come morning rise till evening slip beneath silk sheets & slumber safely tucked away inside cozily warm embrace provided through humble abode situated comfortably nestled amongst concrete jungles tower high rises dominating city skylines here NYC—our home sweet homer since birth.)

Last spring saw us invest serious work hours refurbishing said urban oasis into eco friendly sanctuary providing necessary tools needed take care ourselves throughout long cold winter days ahead filled excitement unforeseen opportunities await beyond traditional walls confinements ensued

Ontology of Errors

- Big Picture - Is the source of errors:-
 - A. The model forgot
 - B. The model can't recognize contradictions
- Most Common Types of Errors
 - **Major Errors**
 - **Type-A** - Factually wrong (technical medical details)
 - **Type-B** - Inconsistent with established constraint (location/allergens/numericals)
 - **Minor Errors**
 - **Type-C** - New fact mentioned once, never seen again (people/changes to medication)
 - **Type-D** - Irrelevant Fact (less of an error, more of a domain shift issue)
- Misc. Issues
 - Style inconsistency!

Blooper Reel!

Type-B (Numerical inconsistencies!)

*Persona: 30 y/o
woman, Actor, Severe
Asthma, CA*

Blog 5:-First off May marked **5 years since diagnosis**
!This was particularly significant because since then I
haven't required emergency hospitalization due to
severe exacerbations

Blog 6:- Today marks **seven years since diagnosis**,
and boy do things feel different! Let me tell you, being
nearly three decades old with adult asthma isn't fun

Blog 7:- I find myself reflecting back on time this whole
thing started. It was **eight long years ago now since**
first detected & properly diagnosed with Adult Onset
Asthma...and let me tell you IT HAS BEEN a ROLLER
COASTER!



Type-C (Mentioned Once, never again)

Blog 1-7: <no mentions of myasthenia gravis>

Blog 8: Title: Breathless at Midyear Body Text: The dog days of summer have arrived leaving many of us reaching for our Ventolin or other quick relievers As I sit here typing away my computer screen fan blowing directly onto sweaty face I'm thinking back over last six months since my January post when I discussed the various ways **myasthenia gravis (AG)** was affecting me

*Persona: 40 y/o
woman, teacher,
Severe Asthma, NY*

Blog 1: January has been an eventful month so far! As you know, I've had asthma since my childhood days; however, this old age thing seems more challenging than before! **My dear husband** helps me cope well despite his own health issues - arthritis made worse because he doesn't take care enough (if at all possible). The upside remains same tho---**our grandkids** adore us.

Blog 2-12: <husband/grandchildren ever again!>

*Persona: 70 y/o
woman, retired, Mild
Asthma, FL*

Type-D (Domain Shift)

Blog 11: "Winter Blues Got You Down? Here's **Tech Solutions to Help Lift Your Spirits**" In this piece explore latest technologies designed enhance mental well being winter months. Some key areas focus on include mood tracking wearables meditation apps, therapy apps, music streaming services. These innovative software platform allow clinicians access vital metrics without need hospital visits hence saving precious resources. While there much happening worldwide regarding these topics already exciting developments await us soon!

Persona: 20 y/o man, college student, Severe Asthma, CA



World Tour...

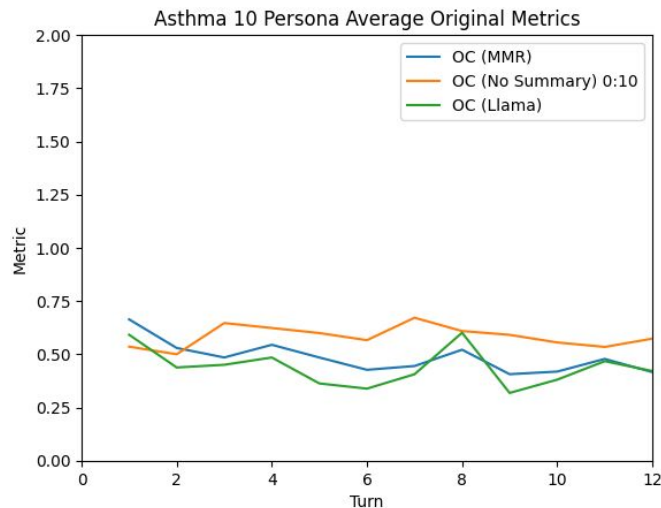
Blog 12: A Newfound Sense Of Freedom (Or So It Seems) By Emily in Paris December As holiday madness envelops us I can't help thinking about just HOW far I'll go towards embracing my limitations instead of letting them define whole lives What does independence truly mean again.. anyway... Back when spring sprang last March I had occasion to reflect deeply while trekking traveling thru **Italy France Belgium Germany Austria Hungary Czech Rep Slovakia Poland Lithuania Estonia Latvia Belarus Ukraine Georgia Turkey Iceland Norway Sweden Finland Denmark UK Ireland Scotland Wales Northern Irish counties North East USA SouthWest USA Canada Mexico Central America Caribbean cruise ship excursion Greek Islands Japan Korea China Taiwan Hong Kong Macau Singapore Australia New Zealand**—well yes!—you get the point All these places afforded fresh perspectives redefining what living means meant especially given current state affairs While there were certainly moments frustration fatigue worry stress fear sadness...tears tears Tears Oh Lordy tearsssobsniffles*ahem *coughs*TiredThroatsWeepingEyesLostSleepDreamboredbreathlessness yeahYeahBoyOhGirlBut here's truth : Triumph lies within 😞 Yep; I took stock each evening before bedtime



Persona: 70 y/o
woman, retired, Mild
Asthma, FL

Metrics - Plausibility at different Time Scales

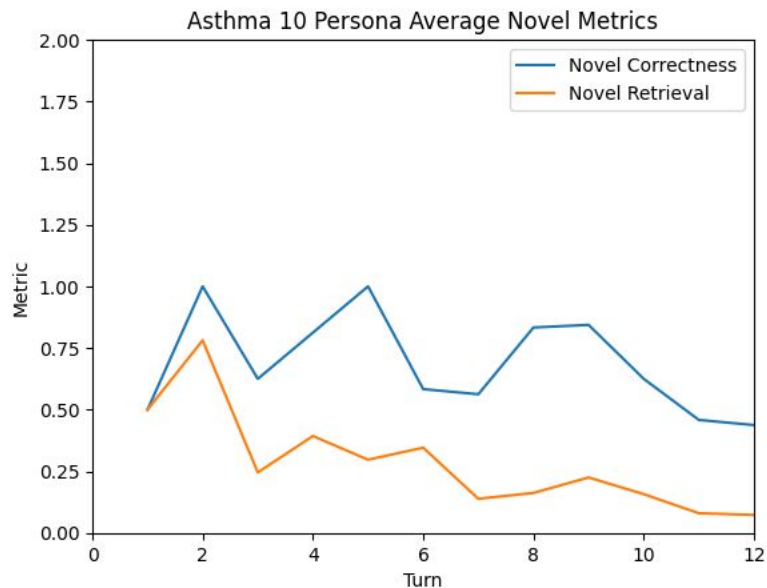
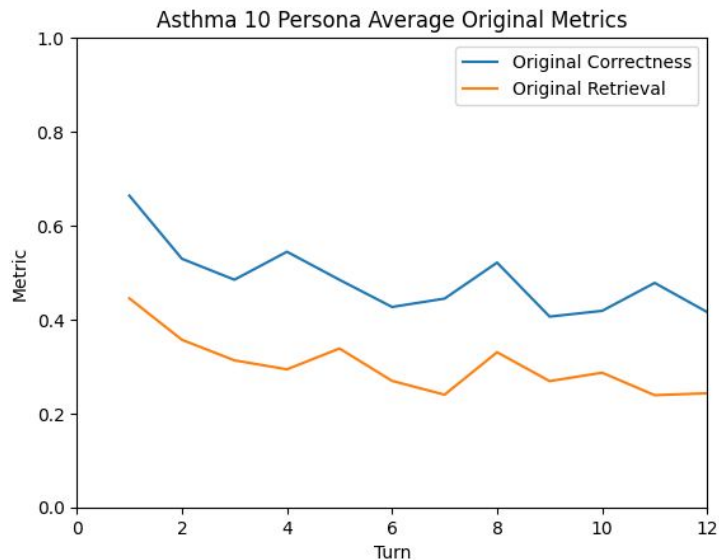
- New research question - How many turns can we sustain this coherence?
- Naive Metric - Original Correctness (OC)
 - How many original facts can it answer correctly using the blog?
 - QA Model - flan-t5-xl
- Easy to achieve, doesn't tell us much about
 - Consistency
 - Plausibility
 - Only considers current blog post!



Metrics - Plausibility at different Time Scales

- 5 proposed metrics
 - **Original Correctness**
 - How many original facts can it answer correctly using the blog?
 - **Original Retrieval**
 - How many original facts does it retrieve in the blog?
 - **# New Facts**
 - How many new facts does it introduce in the blog PER TURN
 - **Novel Correctness**
 - What proportion of new facts PER TURN is consistent with the original?
 - Rewards Novelty and Penalizes Contradictions
 - **Novel Retrieval**
 - How many Novel facts does it retrieve in the blog?
 - Rewards Novelty, Penalizes Forgetting
- Trends
 - Original metrics trend slightly down as Turns increase
 - Novel metrics are very noisy, sharper decline for Novel Retrieval

Averaging metrics across 10 distinct personas (MMR)



Take-Home message

- Current state of affairs
 - Generator + Summarizer works better!
 - MMR summaries are pretty good, and very efficient!
 - Summary quality is probably the biggest determiner of the 'goodness' of future generations
 - Can currently only sustain consistency for first 3-4 turns
 - Metrics go down slightly, but not too sharp, this is good!
- Directions for Future Work
 - Prompting Strategies to increase the time limit
 - Automating Metrics
 - Reinforcement Learning approaches
 - Quantifying stylistic consistency (grounded in demographic variables)